

TITLE: WEB SITE LOAD MANAGEMENT

INVENTORS: DAVID B. KINDER, STANLEY MO and
LINDA B. WELSH

Date: May 19, 2000

WEB SITE LOAD MANAGEMENT

Background

This invention relates generally to Internet web servers.

Internet web servers post web pages that form web
5 sites for various business and other entities. In a number
of circumstances, these web sites may experience excessive
loads. In fact, the number of attempted accesses may
become so high that the web server hosting the web site may
fail. As a result, users may become frustrated with the
10 web site's performance. In cases of commercial web sites,
this may result in lost sales and decreased customer
satisfaction.

Often there is a disconnection between the business
that owns the web site and the web site server. For
15 example, it is conventional that a business may establish a
web site at a web hosting facility. The web hosting
facility may have no way to know anything about the
business activities of the company whose web page it hosts.

As a result, the company may place advertisements or
20 other broadcast information, including its uniform resource
locator (URL), to facilitate access to its web page without
notifying the web hosting facility. In some cases, the

resulting web page access load may be so excessive that the web server may fail.

As an example of circumstances wherein a web server failure may occur, a URL may be provided in a widely broadcast television advertisement. A large number of viewers, seeing the advertisement at the same time, may attempt to simultaneously access the web site. If the number of attempted accesses is excessive, the web server may fail, making it impossible for a large number of potential customers to access the desired information. This obviously results in lost business opportunity and wasted expense for television advertising that does nothing but frustrate potential customers and decrease goodwill.

The problem may be even more acute in systems which broadcast television programming together with enhanced content or ancillary data. In these systems, URLs may be broadcast with the television programming. For example, users may receive these broadcasts on set-top boxes that enable them to mouse click on the information provided with the programming to immediately access the associated web site. Thus, it is particularly easy for set-top box users to quickly access associated web pages broadcast with television programming. As a result, the possibility of web site server failure is increased because many viewers receive the television content together with the ancillary data at the same time. As a result, they may all attempt

to access the associated web site at the same time. This may result in prolonged server failure with extended downtime.

Thus, there is a need for better ways to manage attempted web accesses in periods of high interest.

Brief Description of the Drawings

Figure 1 is a schematic depiction of one embodiment of the present invention;

Figure 2 is a flow chart for software in accordance with one embodiment of the invention which may be stored on the scheduler shown in Figure 1; and

Figure 3 is a flow chart for software in accordance with one embodiment of the present invention that may be stored at the web hosting facility shown in Figure 1.

Detailed Description

A web server management system 10, shown in Figure 1, includes a scheduler 16 which schedules the transmission of television programming and associated enhancements such as ancillary information. The scheduler 16 may be in accordance with the Advanced Television Enhancement Forum (ATVEF) Specification, Version 1.1r.26 (2/2/99) for example. Television programming may be transmitted together with ancillary data such as uniform resource locators (URLs). The programming and related ancillary data may be programmed for transmission at the same time using the so-called

transport A protocol. Alternatively, the ancillary data may be transmitted at different times from the programming information for eventual re-linking with the programming information, under the transport B protocol.

5 The scheduler 16 controls the schedule of transmission of ancillary data 14 received from content providers. The scheduler 16 transports the content 14 over a transport 18 to a plurality of receivers 24.

10 The receivers 24 may be television receivers or systems with information handling capability which may be known as processor-based systems. These processor-based systems may provide Internet access in addition to simply receiving television programming. One such receiver is the so-called 'set-top box that uses a television receiver as
15 the display and implements conventional desktop computer functionality including Internet access. Thus, a plurality of receivers, such as the receiver 24, may be coupled to both the transport 18 and the Internet 22.

20 The transport 18 may also communicate the same programming information and ancillary data to a web hosting facility 26. Thus, the web hosting facility 26 may receive the television programming information. This information may be of use to the web hosting facility as a last warning that ancillary data, including a URL, has been broadcast to
25 a large number of receivers. This may also warn the

facility 26 that a large number of attempted accesses to a URL, hosted by the facility 26, may be imminent.

The scheduler 16 may also develop control code 12.

The control code 12 provides information about URLs that

5 will be broadcast as ancillary data with programming.

Thus, the control code 12 may include information about the URL, the channel on which it is being broadcast, the time or times when it is broadcast and the geographical area to which it is being broadcast. This information may be

10 included within an announcement stream in a digital broadcast as additional ancillary data. The control code 12 may also be included in the vertical blanking interval (VBI) or closed caption stream in connection with a conventional analog broadcast.

15 The control code 12 may be sent out for parsing by the web hosting facility 26 that may receive it through the transport medium 18. The facility 26 may also receive the same warning earlier, directly from the scheduler 16. That is, at the time the scheduler 16 sets the schedule for
20 transmitting the ancillary data 14, the scheduler 16 may provide an early warning to the web hosting facility 26 of the intended schedule of the planned URL announcement.

This scheduling information may, for example, be transmitted over the Internet 22 through a back channel 20
25 from the scheduler 16 to web hosting facility 26 over the Internet 22.

Knowing when the potential load spike may occur, the web hosting facility 26 may schedule extra hosting capacity. For example, the web hosting facility 26 may schedule a surge site 28 to handle a particular URL. That is, a very high capacity surge site 28, that may include a large number of servers, may be assigned the task of handling incoming accesses to a given URL. As a result, a very large load of attempted accesses may be efficiently handled without server failure.

This surge site 28 may be a shared resource for hosting surge pages for multiple non-concurrent broadcast events. The surge site 28 may thereby achieve an economy of resources, conserving the capabilities of a plurality of servers.

Referring to Figure 2, the software 30 may be stored in the appropriate memory within the scheduler 16 in one embodiment of the present invention. The scheduler 16 is itself an information handling system or processor-based system, such as a server, in accordance with one embodiment of the present invention. Thus, the software 30 may be stored in a memory associated with the server 16 for execution by a processor included within the server 16.

The scheduler software 30 initially checks to determine whether ancillary data has been received as indicated in diamond 32. This data may be received from various content providers for distribution over a transport

18. The transport 18 may be any of a variety of conventional transports for providing television content including airwave broadcasts systems, satellite transmission systems or cable distribution systems.

5 Once the information has been received, the software 30 schedules the transmission of that information as indicated in block 34. Further, the control code 12 is developed as indicated in block 36. For example, the control code 12 may provide information about the URL that
10 may be broadcast as well as its broadcast times, geographical area and other pertinent information.

 In addition, the scheduler software 30 may immediately notify the web hosting facility 26 affected by a given URL broadcast, as indicated in block 38. Once the URLs that
15 are going to be broadcast with television programming are identified, a lookup table may be utilized to identify the web hosting facility 26 associated with each broadcast URL. As a result, the web hosting facility 26 may be automatically notified, for example, over the Internet 22
20 using a back channel 20. Should the schedule be changed, a scheduler may send an updated notification of a new surge time.

 A check at diamond 40 determines whether the transmission time for any content to be transmitted has
25 arrived. If so, that content is transmitted as indicated

in block 42. The transmitted content may include programming, ancillary data and the control codes.

Turning next to Figure 3, the software 44 may be stored in a memory in a processor-based system, such as a server at the facility 26, in one embodiment of the invention. Initially, a check at diamond 46 determines whether the control code 12 has been received. It may be received through a back channel 20 or over the transport 18 that also provides the programming information. The control code 12 may be parsed from the programming information. Once the control code is identified, it may provide sufficient information to determine what the nature of the broadcast will be and to enable the facility 26 to predict its load at given periods of time. Thus, the facility 26 assesses the load as indicated in block 48.

A check at diamond 50 determines whether the expected load exceeds the available capacity. If so, a surge site may be scheduled as indicated in block 52. Again, the surge site may be a very high capacity site that can handle a large number of accesses over a relatively short period. The surge site may be a shared resource for non-concurrent surge events.

While the present invention has been described with respect to a limited number of embodiments, those skilled in the art will appreciate numerous modifications and variations therefrom. It is intended that the appended

